الفصل الأول

المقدمة

1.نشوء وتطور علم الإحصاء

علم الإحصاء: علم قديم كقدم المجتمع البشري حيث ارتبط منذ نشأته بعمليات العد التي كانت تجريها الدولة في العصور الوسطى لحساب أعداد جيوشها والضرائب

التي تجبى من المزارعين وجمع المعلومات عن الأراضي التي تسيطر عليها الدولة وغيرها.

أصل كلمة الإحصاء statistics مشتقة إما

1.من اللاتينية status

2.من الايطالية statista

3.من الألمانية statistik

وكلها اشتقت من كلمة دولة state

ثم التكلم عن ظهور الأعداد العربية 4,.....

ثم التكلم عن الحروف الهندية 1,2,3,4,,,,,,,,,,,,,,,,,,,,,,,,,,,,

ثم نبذة تاريخية عن تطور علم الإحصاء منذ القرن السابع عشر حيث شهد ولادة الإحصاء الحيوي vital statistic وكذلك تكونت أساسيات نظرية الاحتمالات probability theory والتي تعتبر العمود الفقري لعلم الإحصاء ثم نظرية المباريات game theory

<u>تعريف علم الإحصاء:</u>

1. عرفه bodding ton بأنه علم التقديرات والاحتمالات

2. عرفَه Lovitt بأنه العلم الذي يختص بجمع وتصنيف وتبويب الحقائق العددية كأساس لتفسير ووصف ومقارنة الظواهر

3.عرفَه cowden croxton بأنه العلم الذي يختص بجمع وتحليل وتفسير البيانات العددية

4. أما المشهداني فعرفه بأنه الطريقة العلمية التي تختص بجمع البيانات والحقائق بالشكل الذي يسهل عملية تحليلها وتفسيرها ومن ثم استخلاص النتائج واتخاذ القرار على ضوء ذلك

والإحصاء يتكون من فرعين هما:

1.الإحصاء الوصفي :

والذي يهتم بـــ:

ا. جمع البيانات

ب.تنظيمها وتوصيفها وتبويبها

ت.أمكانية عرضها في جداول ورسومات بيانية

ث.حساب بعض المؤثرات الإحصائية (النزعة المركزية والتشتت وغيرها)

الطريقة الإحصائية في البحث العلمي: يعني توفر البيانات والمعلومات عن الظاهرة المطلوب دراستها

أي أمكانية التعبير عن الظاهرة تعبير كميا لذا فان هذه الطريقة تستخدم من قبل الباحثين في جميع الفروع الأخرى كونها تهيئ أسلوب علمي وموضوعي محايد دون تمييز أو تدخل من قبل الباحث

أما المراحل الرئيسية للطريقة الإحصائية في البحث العلمي فهي

1.تحديد المشكلة

2.جمع البيانات عن الظاهرة

3.تصنيف وتبويب البيانات وعرضها

4.حساب المؤثرات الإحصائية

5.تحليل

6.تفسير النتائج

اسلوب تصميم البحوث

يجب أن يراعي الباحث مايلي عند تصميم بحثه:-

1.تحديد هدف البحث بشكل واضح ودقيق

2.تحديد إمكانية التنفيذ الفعلي للبحث / المتطلبات المادية والبشرية

3.تحديد إطار البحث (أي تحديد المجتمع الإحصائي)

4.تحديد اسلوب جمع البيانات والمعلومات

هل يستخدم اسلوب التسجيل الشامل أو العينات

الفصل الثاني

جمع وتصنيف وتبويب البيانات:

- **مصادر جمع البيانات:هناك مصدران لجمع البيانات يعتمدها الباحث هي:**

1.المصادر التاريخية:هي كل المعلومات المتجمعة لدى أجهزة ومؤسسات ودوائر الدولة والمحفوظة لديها لسنوات سابقة مثل بيانات التعداد العام للسكان , إحصاءات التجارة الخارجية وغيرها

- **2.مصادر الميدان:هي كل المعلومات التي يجمعها الباحث في حينها من مصادرها الأصلية بأي وسيلة كانت سواء مراسلة أو مقابلة أو أخرى مثل وقوعات الزواج والطلاق خلال سنة معينة .**

أساليب جمع البيانات

1.أسلوب التسجيل الشامل:و هي جمع بيانات عن جميع المفردات التي تؤلف المجتمع الإحصائي هنا يجب أن يكون المجتمع محدد

2.أسلوب العينات:وهي جمع البيانات والمعلومات عن جزء من المفردات التي تؤلف المجتمع الإحصائي.وهو أسلوب مفيد في المجتمعات الغير محدودة كما وتحتاج إلى وقت وجهد وموارد مادية وبشرية اقل مما يحتاجه أسلوب التسجيل الشامل.

وتقسم العينات إلى قسمين هما:

1.العشوائية:هنا لا يكون للباحث دخل أو رأي في اختيار هذه المفردة أو تلك وإنما على أساس عشوائي وتكون على أربعة أصناف هي:

ا. بسيطة:شرط سحبها أن يكون المجتمع الإحصائي متجانس وكل مفردة في المجتمع لها نفس الفرصة في الظهور ضمن مفردات العينة

ب-الطبقية:شرط سحبها أن يكون المجتمع غير متجانس بحيث نستطيع تحويله إلى مجتمع متجانس عن طريق تقسيمه إلى L طبقة بحيث نضمن مبدأ التجانس داخل كل طبقة ويتم سحب عينة عشوائية بسيطة من كل طبقة بحيث يتناسب حجمها مع حجم الطبقة في المجتمع

ج-منتظمة:شرط سحبها هو أن يكون المجتمع الإحصائي مرتبا ترتيبا تصاعديا أو تنازليا أو وفق أي ترتيب أخر حيث نقسم المجتمع المرتب إلى L طبقة ثم يتم سحب عينة عشوائية بسيطة من الطبقة الأولى فقط ثم نضيف K لنسحب المفردة الثانية وهكذا بحيث نضمن سحب مفردة من كل طبقة.

د-متعددة المراحل:شرط سحبها هو أن يكون المجتمع الإحصائي كبير جدا وموزع على بقعة جغرافية واسعة كأن يكون المجتمع على مستوى البلد أو المحافظة أو المقاطعة أو إقليم هنا نبدأ بتقسيم المجتمع إلى وحدات أولية ثم نأخذ عينة عشوائية من هذه الوحدات كمرحلة أولى. ثم نقسم كل وحدة أولية مختارة إلى وحدات اصغر (الوحدات الثانوية) ويتم اختيار عينة من الوحدات الثانوية لكل وحدة أولية كمرحلة

ثانية. ونستمر بالتقسيم والاختيار على هذا المنوال إلى إن نصل إلى عدد المفردات التي تؤلف العينة العشوائية

2.العينات الغير عشوائية: هنا يتدخل الباحث شخصيا في اختيار مفردات العينة وبصورة متحيزة وليس على أساس عشوائي لاعتقاده بأن هذه المفردات هي خير ما تمثل المجتمع والتي ستغني الدراسة بآرائها وأطروحاتها و أهم أنواعها هي الحصصية والعمدية .

الأخطاء الشائعة في جمع البيانات:

هناك بعض الأخطاء التي تحدث عندما يقوم الباحث بجمع البيانات والمعلومات التي تخص بحثه :

1-خطأ التحيز :هنا هذا الخطأ يرتكبه المصدر أو المفردة الإحصائية التي تزود الباحث بالمعلومات سواءا بقصد أو بغير قصد أو يحدث هذا الخطأ أحيانا عندما يستقي الباحث معلومات بحثه ليست من مصادرها الأصلية بل من مصادرها الغير مباشرة

2-خطأ الصدفة: هذا الخطأ يرتكبه الباحث بنفسه سواءا بتعمد أو بصورة غير متعمدة حيث يستقي معلومات بحثه بالاعتماد على ذاكرته بسبب بعد المفردة

الإحصائية عنه أو لأي سبب شخصي أخر .هذا سيؤدي إلى الحصول على نتائج واستنتاجات غير دقيقة وبعيدة عن الواقع.

تنظيم وتصنيف وتبويب البيانات

ان البيانات التي تم جمعها والتي تخص مجتمع البحث غالبا ما تكون بيانات أولية غير مصنفة وغير منتظمة بحيث يتعذر على الباحث تكوين فكرة عن المشكلة قيد البحث وبالتالي لا يمكن إجراء التحليل الإحصائي للوصول الى نتائج البحث 0لذا يتطلب الأمر مراجعة وتدقيق البيانات التي تم جمعها للتأكد من تكاملها ودقتها ووضوحها واستبعاد الاستمارات التي تحوي معلومات ناقصة أو غير دقيقة

تصنيف البيانات

بعد أن يتأكد الباحث من دقة وتكامل البيانات التي تم جمعها تأتي عملية تصنيف البيانات وتعني تشكيل مجموعات على أساس صفة معينة كان تكون ظاهرة الطول أو الوزن أو العمر أو الجنس أو الإقامة وغيرها

تبويب البيانات

وتعني تفريغ البيانات المصنفة وترتيبها في جداول خاصة بحيث إن كل جمع منها يخص مستوى معين 0

وهناك أربعة أشكال للتبويب هي الزمني؛الجغرافي؛الكمي؛والتبويب على أساس صفة معينة(أي أن نوع المتغير وصفي)0ومن الجدير بالذكر أن الغرض من عملية التبويب هو من اجل إبراز وعرض البيانات في أضيق حيز بحيث يتمكن الناظر إليها من تكوين فكرة عنها بوقت قصير وبدون عناء

الفصل الثالث

التوزيعات التكرارية وأساليب عرض البيانات

المتغيرات العشوائية:

المتغير العشوائي:هو دالة ذات قيمة حقيقية معرفه على فضاء العينة و تقسم المتغيرات الى قسمين:

ا. المتغيرات النوعية qualitative variables

هي المتغيرات التي لا يمكن قياسها بوسائل قياس مألوفة وإنما تشكل صفات لذلك المتغير مثل لون العين كمتغير (سوداء؛جوزي؛خضراء؛ زرقاء000),الجنس كمتغير(ذكر؛أنثى),الحالة الاجتماعية كمتغير(أعزب؛ متزوج؛مطلق؛أرمل)

ب- المتغيرات الكمية quantitative variables

هي المتغيرات التي يمكن قياسها بوسائل قياس مألوفة مثل عدد المرضى الراقدين في مستشفى

بعقوبة العام؛عدد رؤوس الماشية في قطيع معين؛درجات الطلبة في كلية؛أطوال الأشخاص بال

بالسنتمترات؛أوزان الأشخاص بالكيلوغرامات؛ ودرجات الحرارة في مدينة معينة0

وتقسم المتغيرات الكمية الى قسمين هما:-

ا-المتغيرات المتقطعة(المتصلة) discrete variables

اذاكانت مجموعة القيم التي ياخذها المتغير العشوائي مجموعة قابلة للعد أي نستطيع من عدها سواء اكانت مجموعة محدودة او غبر محدودة مثل عدد اشجار النخيل في قرية معينة؛عدد طلبة الصف الاول في مدارس بعقوبة وغيرها000000

ب-المتغيرات المستمرة (المتصلة) continuous variables

اذا كانت مجموعة القيم التي ياخذها المتغير العشوائي مجموعة غير قابلة للعد سواء كانت محدودة او غير محدودة وانما تشكل قيم واقعة ضمن فترات0 وهذا يعني وجود عدد غير منته من القيم مثل كمية الامطار المتساقطة على منطقة خلال سنة معينة؛اسعار سلعة معينة في فترة زمنية في فترة زمنية معينة وغيرها000

التوزيعات التكراريـــــــــة

التوزيع التكراري:هو بيانات المتغير العشوائي المجمعة والمصنفة الى عدد من المجاميع تسمى ب((الفئات))

هذه الفئات تكون غالبا مرتبة ترتيبا تصاعديا او حسب طبيعة البيانات0

بكلمات اخرى التوزيع التكراري هو توزيع قيم x حسب الفئات0 ويفضل ان تكون اطوال الفئات متساوية في الطول لكي تسهل العمليات الحسابية

كيفية تكوين جدول توزيع تكراري

لاجل تكوين جدول توزيع تكراري نتتبع الخطوات التالية:-

1- ايجاد المدى الكلي للتوزيع الذي يمثل الفرق بين اكبرقيمة واصغر قيمة مضافا لها العدد واحد

2- ايجاد عدد الفئات حسب صيغة ستيور جيس

$$M= 1+3.322\log n$$

3- ايجاد طول الفئة L حيث ان $L=R\backslash M$

وافضل توزيع تكراري يجب ان تكون

ا. عدد الفئات متساوي الطول

ب.ان يكون التوزيع مغلق وذلك لتسهيل العمليات الحسابية ولكي نتمكن من حساب بعض مقايس النزعة المركزية مثل الوسط الحسابي وبعض مقاييس التشتت مثل الانحراف المعياري والتباين وغيرها لان حسابها يعتمد على تحديد مراكز الفئات التي لايمكن من تحديدها اذا كان الجدول مفتوحا

ت.حدود الفئات بدايتها ونهايتها يجب ان تكون اعداد صحيحة لتسهيل العمليات الحسابية

ث.عدد الفئات يتراوح بين (5-15)

التوزيع التكراري المتجمع: هو كمية التكرار المتجمع عند نقطة معينة من قيم المتغير العشوائي ويكون على نوعين هما

1-التوزيع التكراري المتجمع الصاعد :

هو تراكم التكرارات ابتداء من الفئة الاولى وانتهاء بالفئة الاخير بالاعتماد على

ا. الحدود العليا للفئات

ب.عبارة اقل من اقل من او يساوي

2 التوزيع التكراري المتجمع النازل:

هو تناقص التكرارات ابتداء من الفئة الاولى وانتهاء بالفئة الاخيرة وبالاعتماد على

ا- الحدود الدنيا للفئات

عبارة اكبر من او يساوي

العرض الهندسي للبيانات

تستخدم لغرض اعطاء فكرة واضحة وسريعة عن البيانات وهذه الاشكال هي

1. **الاشرطة البيانية**
2. **الدائرة البيانية**
3. **المستطيل البياني**
4. **الخطوط البيانية**
5. **المربع التكراري**
6. **المضلع التكراري**
7. **المنحنى التكراري**